

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:22:47 ; Search time 18.74 seconds
(without alignments)
243.889 Million cell updates/sec

Title: US-09-619-198-7

Perfect score: 292

Sequence: 1 QAEATRQAAQAERLADLAS.....GLQETQQRNREERAEQ 60

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR-68:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|--------|-------------------------|
| 1 | 292 | 100.0 | 617 | 2 | I56530 | gene VGF protein - rat |
| 2 | 156 | 53.4 | 711 | 2 | S05381 | VGF8a protein precursor |
| 3 | 75.5 | 25.9 | 568 | 2 | S15008 | gene disco protein |
| 4 | 73.5 | 25.2 | 622 | 2 | S71342 | calnexin precursor |
| 5 | 70 | 24.0 | 1979 | 2 | C71622 | hypothetical protein |
| 6 | 68 | 23.3 | 966 | 2 | S25365 | CYC8 protein - yeast |
| 7 | 68 | 23.3 | 1353 | 2 | T27404 | hypothetical protein |
| 8 | 67 | 22.9 | 478 | 2 | S56904 | hypothetical protein |
| 9 | 67 | 22.9 | 913 | 2 | T52485 | neurofilament protein |
| 10 | 66.5 | 22.8 | 993 | 2 | T17230 | hypothetical protein |
| 11 | 66.5 | 22.8 | 4687 | 1 | A39638 | plectin - rat |
| 12 | 65.5 | 22.4 | 284 | 1 | TWCHA | tropomyosin 2, ske |
| 13 | 65.5 | 22.4 | 284 | 2 | A28499 | tropomyosin alpha |
| 14 | 65.5 | 22.4 | 284 | 2 | S24402 | alpha-tropomyosin |
| 15 | 65.5 | 22.4 | 284 | 2 | A26113 | tropomyosin alpha |
| 16 | 65.5 | 22.4 | 284 | 2 | S24401 | tropomyosin 2, ske |
| 17 | 65 | 22.3 | 120 | 2 | F72586 | hypothetical protein |
| 18 | 65 | 22.3 | 1023 | 2 | S12519 | glutactin - fruit |
| 19 | 64.5 | 22.1 | 1038 | 2 | T02634 | rep protein homolog |
| 20 | 64 | 21.9 | 92 | 2 | S23527 | embryonic abundant |
| 21 | 64 | 21.9 | 544 | 2 | B44841 | low molecular weight |
| 22 | 64 | 21.9 | 707 | 2 | T01803 | hypothetical protein |
| 23 | 63 | 21.6 | 504 | 2 | S51942 | prunin 2 precursor |
| 24 | 63 | 21.6 | 537 | 2 | A54424 | acrosomal protein |
| 25 | 63 | 21.6 | 678 | 2 | F70404 | flagellar export p |
| 26 | 63 | 21.6 | 1018 | 2 | S44758 | cl4B9.6 protein - |
| 27 | 62.5 | 21.4 | 175 | 2 | S57987 | hypothetical protein |
| 28 | 62.5 | 21.4 | 320 | 2 | S50716 | hypothetical protein |
| 29 | 62.5 | 21.4 | 462 | 2 | I53414 | lamin C2 - mouse |

| | | | | | | |
|----|------|------|------|---|--------|-------------------------|
| 30 | 62.5 | 21.4 | 572 | 1 | VEHULC | lamin C - human |
| 31 | 62.5 | 21.4 | 574 | 2 | S04333 | lamin C - mouse |
| 32 | 62.5 | 21.4 | 664 | 1 | VEHULA | lamin A - human |
| 33 | 62.5 | 21.4 | 665 | 2 | S27267 | lamin A - rat |
| 34 | 62.5 | 21.4 | 665 | 2 | S28182 | lamin A - mouse |
| 35 | 62.5 | 21.4 | 1184 | 2 | G01763 | atrophin-1 - human |
| 36 | 62.5 | 21.4 | 1560 | 2 | T30282 | calcium-binding protein |
| 37 | 62 | 21.2 | 102 | 2 | T01285 | hypothetical protein |
| 38 | 62 | 21.2 | 536 | 1 | A47190 | transducer protein |
| 39 | 62 | 21.2 | 536 | 2 | E84318 | Htr1 transducer [i |
| 40 | 62 | 21.2 | 924 | 2 | S06117 | myosin heavy chain |
| 41 | 62 | 21.2 | 1110 | 2 | I51116 | NF-180 - sea lamp |
| 42 | 62 | 21.2 | 1432 | 2 | B85431 | trichohyalin like |
| 43 | 62 | 21.2 | 2007 | 1 | B43402 | myosin heavy chain |
| 44 | 61.5 | 21.1 | 1087 | 2 | T30330 | gelosolin-related p |
| 45 | 61.5 | 21.1 | 1093 | 2 | A47212 | transcription fact |

ALIGNMENTS

RESULT 1

I56530

gene VGF protein - rat

N:Alternate names: VGF

C:Species: Rattus norvegicus (Norway rat)

C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 01-Dec-2000

C:Accession: I56530; A39748; JH0454

R:Hawley, R.J.; Scheibe, R.J.; Wagner, J.A.

J. Neurosci. 12, 2573-2581, 1992

ATTN: NGF induces the expression of the VGF gene through a cAMP or response element

A:Reference number: I56530; MUID: 92309005

A:Accession: I56530

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-617 <RES>

A:Cross-references: GB:M74223; NID:q207650; PIDN:AAA42336.1; PID:q207651

R:Salton, S.R.J.; Fischberg, D.J.; Dong, K.W.

Mol. Cell. Biol. 11, 2335-2349, 1991

ATTN: Structure of the gene encoding VGF, a nervous system-specific mRNA that is r

A:Reference number: A39748; MUID: 91203852

A:Accession: A39748

A:Molecule type: DNA

A:Residues: 1-175, 'D', 177-617 <SALI>

A:Cross-references: GB:M60522

R:Salton, S.R.J.

J. Neurochem. 57, 991-996, 1991

ATTN: Nucleotide sequence and regulatory studies of VGF, a nervous system-specific

A:Reference number: JH0454; MUID: 91318308

A:Accession: JH0454

A:Molecule type: mRNA

A:Residues: 1-175, 'D', 177-617 <SALI>

A:Cross-references: GB:M60525

A:Experimental source: cultured cell PC12, clone NGF33.1

C:Comment: This protein is induced in PC12 cells to a greater extent by nerve growth

C:Genetics:

A:Gene: VGF

C:Keywords: growth factor

Query Match

Best Local Similarity 100.0%; Score 292; DB 2; Length 617;

Mismatches 0; Conservative 0; Indels 0; Gaps 0;

QY 1 QAEATRQAAQAERLADLASDLLQYLLQGGARQRDLLQGGARQREERAEQ 60

DB 313 QAEATRQAAQAERLADLASDLLQYLLQGGARQRDLLQGGARQREERAEQ 372

RESULT 2

S05381

VGF8a protein precursor - rat

C:Species: Rattus norvegicus (Norway rat)

```

F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-622/Product: calnexin #status predicted <NAT>
F:478-500/Domain: transmembrane #status predicted <TM>

Query Match          25.2%   Score 73.5; DB 2; Length 622;
Best Local Similarity 30.0%; Pred. No. 2.8;
Matches 18; Conservative 10; Mismatches 25; Indels 7; Gaps 1;

Qy 1 QAEATROAAQEEERLADLASDLLQYLQGGARQDGLGGRLQTOQERENEREAEQEE 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 546 EEKAKKEAQEDNAEQAEK-----QTGEGEAGAOGSQEEEEEEEEEEEE 598

RESULT 5
C71622
hypothetical protein PFB0145c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: C71622
R:Garner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.;
Perteira, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743
A:Accession: C71622
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1979 <GAR>
A:Cross-references: GB:AE001375; GB:AE001362; NID:g3845105; PIDN:AAC71819.1; PID:g384
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0145c

Query Match          24.0%   Score 70; DB 2; Length 1979;
Best Local Similarity 44.7%; Pred. No. 21;
Matches 17; Conservative 9; Mismatches 10; Indels 2; Gaps 1;

Qy 23 LLOYLQGGARQDGLGGRLQTOQERENEREAEQEE 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 314 LLEYELEKNQIDLLNK--QEKEKEKEKEKEKE 349

RESULT 6
S25365
CYC8 protein - yeast (Saccharomycetes cerevisiae)
N:Alternate names: glucose repression mediator; protein YBR0908; protein YBR112c; SSN
C:Species: Saccharomycetes cerevisiae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 11-Jan-2000
C:Accession: S25365; S48277; S45980; S25404; A30906; S44692
R:Mannhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann, H.
Yeast 8, 397-408, 1992
A:Title: Molecular analysis of yeast chromosome II between CMD1 and LYS2: the excision
A:Reference number: S25364; MUID:92327848
A:Accession: S25365
A:Molecule type: DNA
A:Residues: 1-966 <MAN>
A:Cross-references: EMBL:X66247; NID:g3548; PIDN:CAA46973.1; PID:g3550
R:Mannhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann, H.
Yeast 10, 1363-1381, 1994
A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
A:Reference number: S48255; MUID:95208357
A:Accession: S48277
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-966 <MAW>
A:Cross-references: EMBL:X78993; NID:g476045; PIDN:CAA55615.1; PID:g476068
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
R:Feldmann, H.; Mannhaupt, G.; Schwarzlöser, C.; Vetter, I.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45927
A:Accession: S45980

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:26:07 ; Search time 24.67 Seconds
(without alignments)
355.749 Million cell updates/sec

Title: US-09-619-198-7

Perfect score: 292

Sequence: 1 QAEATROAAQEEERLADLAS.....GLQETQOERENREBEAEQ 60

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Se-ched: 473505 segs, 146272329 residues

Tc number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*

1: sp.archaea:*

2: sp.bacteria:*

3: sp.fungi:*

4: sp.human:*

5: sp.invertebrate:*

6: sp.mammal:*

7: sp.mhc:*

8: sp.organelle:*

9: sp.phage:*

10: sp.plant:*

11: sp.rodent:*

12: sp.virus:*

13: sp.vvertebrate:*

14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|-----------------------|
| 1 | 249.5 | 85.4 | 615 | 4 Q9UDW8 | Q9udw8 homo sapien |
| 2 | 249.5 | 85.4 | 616 | 4 O15240 | O15240 homo sapien |
| 3 | 87.5 | 30.0 | 800 | 5 Q9Y102 | Q9Y102 drosophila |
| 4 | 75 | 25.7 | 287 | 12 Q9DWH5 | Q9dwh5 rat cytochrome |
| 5 | 75 | 25.7 | 320 | 11 Q9CXK9 | Q9cxk9 mus musculus |
| 6 | 73.5 | 25.2 | 622 | 13 Q98985 | Q98985 rana rugosa |
| 7 | 72 | 24.7 | 897 | 13 Q13098 | Q13098 xenopus lae |
| 8 | 71 | 24.3 | 316 | 11 Q62254 | Q62254 mus musculus |
| 9 | 71 | 24.3 | 540 | 11 Q62253 | Q62253 mus musculus |
| 10 | 70 | 24.0 | 238 | 5 Q9VK98 | Q9vk98 drosophila |
| 11 | 70 | 24.0 | 1979 | 5 Q96133 | Q96133 plasmodium |
| 12 | 69.5 | 23.8 | 1326 | 2 Q9L2C3 | Q9l2c3 streptomyces |
| 13 | 69 | 23.6 | 808 | 5 Q18409 | Q18409 drosophila |
| 14 | 68.5 | 23.5 | 773 | 13 P79805 | P79805 morone saxa |
| 15 | 68.5 | 23.5 | 3469 | 5 Q9U4I2 | Q9u4i2 drosophila |
| 16 | 68.5 | 23.5 | 3502 | 5 Q9V1J9 | Q9v1j9 drosophila |
| 17 | 68.5 | 23.5 | 3616 | 13 Q9W6V0 | Q9w6v0 gallus gall |
| 18 | 68 | 23.3 | 1353 | 5 Q9XW65 | Q9xw65 caenorhabdi |
| 19 | 68 | 23.3 | 1985 | 11 Q9JIS7 | Q9jis7 mus musculus |

Q9vwh5 drosophila
Q9dle0 mus musculus
Q9cwi3 mus musculus
O13099 xenopus lae
Q9ueu4 homo sapien
Q9bhm3 paramacium
Q9ufw3 homo sapien
Q9ues6 homo sapien
Q9y2l1 homo sapien
Q9y2l1 homo sapien
Q9y87 homo sapien
Q9xer1 sorghum bic
Q91968 gallus gall
Q90740 gallus gall
Q9w4e8 drosophila
P91956 lytechinus
Q9h016 homo sapien
Q9ycu9 aeropyrum p
Q9v1j3 drosophila
Q9luw1 arabidopsis
Q9ne92 leishmania
Q9fwy3 lycopersico
Q9dd42 mus musculus
Q94936 homo sapien
Q90983 dictyosteli
Q94654 arabidopsis
Q9gpd7 drosophila

ALIGNMENTS

RESULT 1

Q9UDW8 ID Q9UDW8 PRELIMINARY; PRT; 615 AA.
AC Q9UDW8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE WUGSC:H.DJ0747G18.3 PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=99063792; PubMed=9847074;
RA Sulston J.F., Waterston R.;
RT "Toward a complete human genome sequence."
RL Genome Res. 8:1097-1108(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Cordes M., Doella D.;
RT "The sequence of Homo sapiens PAC clone RP4-747G18.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.H.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004876; AAD45830.1; -. 198097C5622AC087 CRC64;
SQ SEQUENCE 615 AA; 67257 MW; 198097C5622AC087 CRC64;

Query Match 85.4%; Score 249.5; DB 4; Length 615;

Best Local Similarity 88.5%; Pred. No. 5.1e-20;

Matches 54; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 QAEATROAAQEEERLADLASDLLLYLQGGARQDLGGRGLOFTQOERENR-EEBAEQ 59

|||||

Db 310 QAEATROAAQEEERLADLASDLLLYLQGGARQDLGGRGLOFTQOERENR-EEBAEQ 369

Y 60 E 60
b 370 E 370

RESULT 2

D O15240 PRELIMINARY; PRT; 616 AA.

C O15240: (TREMBLrel. 05, Created)

T 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

T 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

E NEURO-ENDOCRINE SPECIFIC PROTEIN VGF.

N VGF.

S Homo sapiens (Human).

C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

X NCBI_TaxID=9606;

N [1]

F SEQUENCE FROM N.A.

C ISSUE-PLACENTA;

X MEDLINE=98008940; PubMed=9344675;

A Canu N., Possenti R., Ricco A.S., Rocchi M., Levi A.;

T Cloning, structural organization analysis, and chromosomal assignment

T of the human gene for the neurosecretory protein VGF.*;

L Genomics 45:443-446(1997).

R EMBL=V12661; CAA73210.1

Q SEQUENCE 616 AA; 67286 MW; CD1920610201BEB9 CRC64;

Query Match

Best Local Similarity 85.4%; Score 249.5; DB 4; Length 616;

Matches 54; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Y 1 QAEATQAAQAERLADSLDLYLGGARQDLGGGQLOETQOERNEER-EDEAEQ 59-

b 310 QAEATQAAQAERLADSLDLYLGGARQDLGGGQLOETQOERNEER-EDEAEQ 369

Y 60 E 60

b 370 E 370

RESULT 3

D OY102 PRELIMINARY; PRT; 800 AA.

C OY102: Q9VP39;

T 01-NOV-1999 (TREMBLrel. 12, Created)

T 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

E BCDNA:GH11973 PROTEIN.

N BCDNA:GH11973 OR CG6014.

S Drosophila melanogaster (Fruit fly).

C Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

C Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

C Ephydroidea; Drosophilidae; Drosophila.

X NCBI_TaxID=7227;

N [1]

P SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).

C STRAIN-BERKELEY;

X MEDLINE=20196006; PubMed=10731132;

A Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

A Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

A George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

A Sutton G.G., Wortman J.R., Fandell M.D., Zhang Q., Chen L.X.,

A Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

A Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

A Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

A Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

A Beeson K.Y., Benos P.V., Berman B.P., Brokstein P., Brottier P.,

A Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

A Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

A

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J.J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

The genome sequence of Drosophila melanogaster.;

Science 287:2185-2195(2000).

[2]

SEQUENCE FROM N.A. (LONG ISOFORM).

RC STRAIN-BERKELEY; TISSUE=HEAD;

RX MEDLINE=20196012; PubMed=10731138;

RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E.,
Stapleton M., Harvey D.A.;

RT "A Drosophila complementary DNA resource.";

RL Science 287:2222-2224(2000).

CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.

DR EMBL; AE003594; AAF51717.1; ;
EMBL; AF145671; AAD38646.1; ;
DR EMBL; AE003594; AAF51718.1; ;
DR FlyBase; FBgn0027542; BCDNA:GH11973.

DR InterPro; IPR001304; lectin_c.

DR Pfam; PF00059; lectin_c; 1.

DR SMART; SM00034; CLECT; 1.

DR PROSITE; PS00041; C.TYPE_LCTIN_2; 1.

KW Alternative splicing

FT VARSPLIC 1 142 MISSING (IN SHORT ISOFORM).

SQ SEQUENCE 800 AA; 97346 MW; D30E9F36CF5D84B7 CRC64;

Query Match 30.0%; Score 87.5; DB 5; Length 800;

Best Local Similarity 38.3%; Pred. No. 0.097;

Matches 23; Conservative 12; Mismatches 16; Indels 9; Gaps 2;

OY 1 QAEATQAAQAERLADSLDLYLGGARQDLGGGQLOETQOERNEER-EDEAEQ 60

Db 502 QAEADQROHADEE-----AEKRLERI-----QKRELEAQQREREQRKQREQEEDQ 552

RESULT 4

Q9DWH5

ID Q9DWH5 PRELIMINARY; PRT; 287 AA.

AC Q9DWH5;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE PR4.

OS R4.

GN Rat cytomegalovirus (strain Maastricht).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Muromegalovirus.

OX NCBI_TaxID=79700;

RN [1]


```

y 1 QRENTQAAQAERLADSLDLLLYLGGARQRDLGGRGLQETQQER-ENREREAEQ 59
      : || :: || |::| :|| :|| | |::| | |::| | |::| | |::| |
b 585 EEEAKEKEGEKEEKEVAEE-----EGGEGEGEGEG--ETEEDKGEVEKEEAET 635

iy 60 E 60
      |
ib 636 E 636

RESULT      8
j62254      PRELIMINARY;          PRT;    316 AA.
c Q62254:
yt 01-NOV-1996 (TREMBLrel. 01, Created)
yt 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
yt 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
ye ALTERNATIVE SPLICING PRODUCT FOR PROACROSIN-BINDING PROTEIN (SP32).
ys ACBP.
ys Mus musculus (Mouse).
xc Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
x Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
x NCBI_TaxID=10090;
[1]
rp SEQUENCE FROM N.A.
rc STRAIN-DDY; TISSUE-TESTIS;
ra Baba T.;
.*Rodent-specific alternative splicing of RNA for mammalian sperm
proacrosin-binding protein, sp32.*;
rt Submitted (SEP-1993) to the EMBL/GenBank/DDBJ databases.
rr EMBL: D17574; BAA04498.1; -.
rr MGD; MGI:185915; Acrbp.
rw Alternative splicing.
ix SEQUENCE 316 AA; 35767 MW; 8905ABBC1CF9BB66 CRC64;

Query Match      24.3%; Score 71; DB 11; Length 316;
Best Local Similarity 32.5%; Pred. No. 2.6;
Matches 25; Conservative 7; Mismatches 27; Indels 18; Gaps

iy 2 AEATQAAAQ--EERLADLASDLLLYLGGGARQRDLGGRG-----LQ 43
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 156 ATATEHAQFPWPERLNNNVVELLSGLGKKDQSSRRPGQEHHKHLEEAQ 215

iy 44 ETQORENREEEAEQE 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 216 EQEQEEEEEEAKOE 232

}
t 9
}
j Q62253      PRELIMINARY;          PRT;    540 AA.
c Q62253:
yt 01-NOV-1996 (TREMBLrel. 01, Created)
yt 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
yt 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
ye PROACROSIN-BINDING PROTEIN (SP32).
ys ACBP.
ys Mus musculus (Mouse).
xc Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
xc Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
x NCBI_TaxID=10090;
[1]
rp SEQUENCE FROM N.A.
rc STRAIN-DDY; TISSUE-TESTIS;
ra Baba T.;
.*Rodent-specific alternative splicing of RNA for mammalian sperm
proacrosin-binding protein, sp32.*;
rt Submitted (SEP-1993) to the EMBL/GenBank/DDBJ databases.
rr EMBL: D17573; BAA04498.1; -.
rr MGD; MGI:185915; Acrbp.
rr InterPro: IPR002350; Kazal.
rr InterPro: IPR001005; Myb_DNA_bind.
rr SMART: SM00280; KAZAL; 1.

```

DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
SQ SEQUENCE 540 AA; 61107 MW; 7FD6BB2E711C0AFF CRC64;

Query Match 24.3%; Score 71; DB 11; Length 540;
Best Local Similarity 32.5%; Pred. No. 4.6;
Matches 25; Conservative 7; Mismatches 27; Indels 18; Gaps

OY 2 AENTQNAAO--BERLADLASDLLLYLQGGARQRDLGGR-----IQ 43
| | | | | : | | | | | : | | | |
Db 156 ATATEHAQFQPWERLNNNVEELLQSLSLGRRDQSSRRPQGQKHQEQIOEHKLEAQ 215
| | | | | : | | | | | : | | | |

OY 44 ETQOERENEEEAQE 60
| : : : | | | | | : | | | | |
Db 216 EQBEQEEEEEAAKE 232

RESULT 10
Q9VK98 PRELIMINARY; PRT; 238 AA.

ID Q9VK98
AC Q9VK98;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ARET PROTEIN.
OS ARET OR CG6319.
GN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.R., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.E., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).

```

DR EMBL; AE003635; AAF53180.2; -.
DR HSP; P09651; IUPI.
DR FlyBase; FBgn0000114; aret.
SQ SEQUENCE 238 AA; 28069 MW; FBEA07BABC2D1805 CRC64;

Query Match          24.0%; Score 70; DB 5; Length 238;
Best Local Similarity 34.9%; Pred. No. 2.5;
Matches 22; Conservative 13; Mismatches 12; Indels 16; Gaps

QY      6 RRAQAQBERLADLAS-----DLLLVLLGGAGRDGLGRGL---QTQQERENEREAEA 57
DB      149 QQTSSVQQQLSQLAAALQDHDL-----TQQELFKQLATMWEYORERERER 200
               :|:::|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      58 EQE 60
DB      201 ERE 203
               ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

P       T    11
O       ID   3
I       O     .
D       O96133 PRELIMINARY; PRT; 1979 AA.
AC       AC   O96133;
DT       DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT       DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT       DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE       DE   HYPOTHETICAL 237.7 KDA PROTEIN.
GN       GN   PFB0145C.
OS       OS   Plasmodium falciparum.
OC       OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX       OX   NCBI_TaxID=5833;
RN       RN   [1]
RP       RP   MEDLINE=99021743; PubMed=9804551;
RX       RX   Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA       RA   Koonin E.V., Shallow S., Mason T., Yu K., Fujii C., Pederson J.,
RA       RA   Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA       RA   Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA       RA   Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT       RT   "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT       RT   falciparum.";
RL       RL   Science 282:1126-1132(1998).
RM       RM   EMBL; AE001375; AAC71819.1; -.
RW       RW   Hypothetical protein.
SQ       SQ   SEQUENCE 1979 AA; 237745 MW; 5C6CD6307AFF7D37 CRC64;

Query Match          24.0%; Score 70; DB 5; Length 1979;
Best Local Similarity 44.7%; Pred. No. 23;
Matches 17; Conservative 9; Mismatches 10; Indels 2; Gaps

QY      23 LLVYLGGAGQRDLGGRGLQETOQERENEREEAEQE 60
DB      314 LLEYELEKKNQIDLNNK--QEKEKEKEKEKEKEKE 349
               ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
               :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 12
Q9L2C3
ID       ID   Q9L2C3 PRELIMINARY; PRT; 1326 AA.
AC       AC   Q9L2C3;
DT       DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT       DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT       DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE       DE   LARGE ALA/GLU-RICH PROTEIN.
GN       GN   SC8F4.01C.
OS       OS   Streptomyces coelicolor.
OC       OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC       OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX       OX   NCBI_TaxID=1902;
RN       RN   [1]
RP       RP   SEQUENCE FROM N.A.
RC       RC   STRAIN=A3(2);
RA       RA   Seeger K.J., Harris D.;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:20:47 ; Search time 12.6 Seconds
(without alignments)
107.159 Million cell updates/sec

Title: US-09-619-198-7

Perfect score: 292

Sequence: 1 QAEATRQAAQAEERLADLAS.....GLQETQOENREEREAQE 60

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

1 number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCrUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 66.5 | 22.8 | 651 | 3 | US-08-650-766-6 |
| 2 | 66.5 | 22.8 | 651 | 3 | US-08-922-635-5 |
| 3 | 66.5 | 22.8 | 1070 | 3 | US-08-922-635-22 |
| 4 | 62.5 | 21.4 | 515 | 3 | US-08-705-660-46 |
| 5 | 62.5 | 21.4 | 515 | 3 | US-08-989-045-46 |
| 6 | 62.5 | 21.4 | 1185 | 4 | US-09-041-886-23 |
| 7 | 61.5 | 21.1 | 414 | 5 | PCT-US93-03077-3 |
| 8 | 61.5 | 21.1 | 1093 | 5 | PCT-US93-03077-1 |
| 9 | 61 | 20.9 | 903 | 2 | US-08-853-310-2 |
| 10 | 59 | 20.2 | 240 | 2 | US-08-114-555A-8 |
| 11 | 59 | 20.2 | 240 | 3 | US-08-559-397A-14 |
| 12 | 58 | 19.9 | 1162 | 2 | US-08-728-323A-2 |
| 13 | 57.5 | 19.7 | 450 | 2 | US-08-665-037-2 |
| 14 | 57.5 | 19.7 | 450 | 2 | US-08-666-067-2 |
| 15 | 57.5 | 19.7 | 450 | 2 | US-08-732-870-2 |
| 16 | 57 | 19.5 | 1958 | 1 | US-07-945-283-2 |
| 17 | 56.5 | 19.3 | 76 | 1 | US-08-253-155A-28 |
| 18 | 56.5 | 19.3 | 77 | 1 | US-08-253-155A-47 |
| 19 | 56.5 | 19.3 | 83 | 1 | US-08-253-155A-38 |
| 20 | 56.5 | 19.3 | 240 | 2 | US-08-114-555A-6 |
| 21 | 56.5 | 19.3 | 240 | 3 | US-08-559-397A-12 |
| 22 | 56.5 | 19.3 | 521 | 4 | US-09-370-368-4 |
| 23 | 56.5 | 19.3 | 3111 | 2 | US-08-460-309-4 |
| 24 | 56.5 | 19.3 | 3111 | 2 | US-08-125-077-4 |
| 25 | 56 | 19.2 | 234 | 4 | US-09-040-483-5 |
| 26 | 56 | 19.2 | 905 | 2 | US-08-574-959A-9 |
| 27 | 56 | 19.2 | 905 | 4 | US-09-357-014-9 |

| | | | | | | |
|----|------|------|------|---|-------------------|--------------------|
| 28 | 56 | 19.2 | 955 | 1 | US-08-006-676B-1 | Sequence 1, Appli |
| 29 | 56 | 19.2 | 955 | 1 | US-08-282-845-2 | Sequence 2, Appli |
| 30 | 56 | 19.2 | 955 | 2 | US-08-428-414A-3 | Sequence 3, Appli |
| 31 | 56 | 19.2 | 955 | 5 | PCT-US94-00324-1 | Sequence 1, Appli |
| 32 | 56 | 19.2 | 1135 | 2 | US-08-574-959A-7 | Sequence 7, Appli |
| 33 | 56 | 19.2 | 1135 | 4 | US-09-357-014-7 | Sequence 7, Appli |
| 34 | 55.5 | 19.0 | 316 | 4 | US-08-098-327E-31 | Sequence 31, Appli |
| 35 | 55 | 18.8 | 162 | 1 | US-08-244-557-2 | Sequence 2, Appli |
| 36 | 55 | 18.8 | 304 | 4 | US-09-042-225-7 | Sequence 7, Appli |
| 37 | 54.5 | 18.7 | 55 | 1 | US-08-199-776-4 | Sequence 4, Appli |
| 38 | 54.5 | 18.7 | 55 | 3 | US-08-663-731-4 | Sequence 4, Appli |
| 39 | 54.5 | 18.7 | 55 | 3 | US-08-879-338-4 | Sequence 4, Appli |
| 40 | 54.5 | 18.7 | 55 | 5 | PCT-US95-02044-4 | Sequence 4, Appli |
| 41 | 54.5 | 18.7 | 1178 | 1 | US-08-199-776-2 | Sequence 2, Appli |
| 42 | 54.5 | 18.7 | 1178 | 3 | US-08-663-731-2 | Sequence 2, Appli |
| 43 | 54.5 | 18.7 | 1178 | 3 | US-08-879-338-2 | Sequence 2, Appli |
| 44 | 54.5 | 18.7 | 1178 | 5 | PCT-US95-02044-2 | Sequence 2, Appli |
| 45 | 54 | 18.5 | 161 | 5 | PCT-US95-05741-11 | Sequence 11, Appli |

ALIGNMENTS

RESULT 1

US-08-650-766-6

; Sequence 6, Application US/08650766D

; Patent No. 6015690

; GENERAL INFORMATION:

; APPLICANT: PILETZ, John E.

; APPLICANT: IVANOV, Tina R.

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A HUMAN IMIDAZOLINE RECEPTOR AND
; TITLE OF INVENTION: METHOD FOR CLONING THE SAME

; FILE REFERENCE: Corrected Sequence Listing

; Patent No. 6015690

; CURRENT APPLICATION NUMBER: US/08/650.766D

; CURRENT FILING DATE: 1996-05-20

; EARLIER APPLICATION NUMBER: US 60/012.600

; EARLIER FILING DATE: 1996-03-01

; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6

; LENGTH: 651

; TYPE: PRT

; ORGANISM: Homo sapiens

US-08-650-766-6

Query Match 22.8%; Score 66.5; DB 3; Length 651;

Best Local Similarity 30.6%; Pred. No. 1.6;

Matches 19; Conservative 12; Mismatches 22; Indels 9; Gaps 3;

QY 1 QAEATRQAAQAEERLADLASLLQYLLQG--GARQDLGGGGLQETQOENREEREA 58

Db 220 QGEGEEDDEEEEDVAEN---RYFEMGPPDVEERGGQG-----EEEEEEDEEA 272

QY 59 QE 60

Db 273 EE 274

RESULT 2

US-08-922-635-5

; Sequence 5, Application US/08922635A

; Patent No. 6033871

; GENERAL INFORMATION:

; APPLICANT: PILETZ, John E.

; APPLICANT: IVANOV, Tina R.

; TITLE OF INVENTION: DNA MOLECULES ENCODING IMIDALINE RECEPTIVE POLYPEPTIDES
; TITLE OF INVENTION: AND POLYPEPTIDES ENCODED THEREBY

; FILE REFERENCE: Corrected Sequence Listing

; Patent No. 6033871

; CURRENT APPLICATION NUMBER: US/08/922.635A

; CURRENT FILING DATE: 1997-09-03

```
EARLIER APPLICATION NUMBER: 08/650,766
EARLIER FILING DATE: 1996-05-20
EARLIER APPLICATION NUMBER: 60/012,600
EARLIER FILING DATE: 1996-03-01
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 651
TYPE: PRT
ORGANISM: Homo sapiens
;-08-922-635-5

Query Match      22.8%; Score 66.5; DB 3; Length 651;
Best Local Similarity 30.6%; Pred. No. 1.6;
Matches 19; Conservative 12; Mismatches 22; Indels 9; Gaps 3;

/ 1 QAEATQAAQAEERLADLASDLLLQYLLQG--GARQDLGGGLOETQOERENEREEAE 58
/ 220 QGQGEEDDEEEEDVAEN---RYFEMGPPDVEEEGGGQ-----EEEEEEDEAE 272
/ 59 QE 60
/ 273 EE 274

RESULT 3
3-08-922-635-22
Sequence 22, Application US/08922635A
Patent No. 6033871
GENERAL INFORMATION:
APPLICANT: PILETZ, John E.
APPLICANT: IVANOV, Tina R.
TITLE OF INVENTION: DNA MOLECULES ENCODING IMIDALINE RECEPTIVE POLYPEPTIDES
FILE REFERENCE: Corrected Sequence Listing
Patent No. 6033871
CURRENT APPLICATION NUMBER: US/08/922,635A
CURRENT FILING DATE: 1997-09-03
EARLIER APPLICATION NUMBER: 08/650,766
EARLIER FILING DATE: 1996-05-20
EARLIER APPLICATION NUMBER: 60/012,600
EARLIER FILING DATE: 1996-03-01
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 1070
TYPE: PRT
ORGANISM: Homo sapiens
;-922-635-22

Query Match      22.8%; Score 66.5; DB 3; Length 1070;
Best Local Similarity 30.6%; Pred. No. 3;
Matches 19; Conservative 12; Mismatches 22; Indels 9; Gaps 3;

Y 1 QAEATQAAQAEERLADLASDLLLQYLLQG--GARQDLGGGLOETQOERENEREEAE 58
b 198 QGQGEEDDEEEEDVAEN---RYFEMGPPDVEEEGGGQ-----EEEEEEDEAE 250
Y 59 QE 60
b 251 EE 252

RESULT 4
S-08-705-660-46
Sequence 46, Application US/08705660
Patent No. 5858683
GENERAL INFORMATION:
APPLICANT: KEESEE, SUSAN
APPLICANT: OBAR, ROBERT
APPLICANT: WU, YING-JYE

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
DETECTION OF CERVICAL CANCER
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault
STREET: 125 High St.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,660
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: GREENHALGH, DUNCAN A
REGISTRATION NUMBER: 38,678
REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 515 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-705-660-46

Query Match      21.4%; Score 62.5; DB 2; Length 515;
Best Local Similarity 31.0%; Pred. No. 3.9;
Matches 18; Conservative 9; Mismatches 22; Indels 9; Gaps 1;

QY 2 AEATQAAQAEERLADLASDLLLQYLLQGAGARQDLGGGLOETQOERENEREEAEQ 59
DB 158 SOLQQLAAKEAKLRDLEDSL-----ARERTSRLLAKEKEREMAEARMQOOQ 206

RESULT 5
US-08-989-045-46
Sequence 46, Application US/08989045
Patent No. 6027905
GENERAL INFORMATION:
APPLICANT: KEESEE, SUSAN
APPLICANT: OBAR, ROBERT
APPLICANT: WU, YING-JYE
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
DETECTION OF CERVICAL CANCER
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault
STREET: 125 High St.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,045
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
```